Tensor based generalization of monogenic wavelets for coherent multiscale local phase analysis of color images

Raphaël Soulard, Philippe Carré

TENSOR BASED GENERALIZATION OF MONOGENIC WAVELETS FOR COHERENT MULTISCALE LOCAL PHASE ANALYSIS OF COLOR IMAGES

Raphaël Soulard and Philippe Carré

Xlim-SIC laboratory, University of Poitiers, France

raphael.soulard@univ-poitiers.fr carre@sic.univ-poitiers.fr

ABSTRACT

We propose a new color extension of the monogenic wavelet transform. Monogenic wavelets give a coherent representation of scalar images through a local phase concept and an underlying orientation analysis. We here define a color extension of the monogenic framework. The underlying local geometric analysis and phase concept are generalized by using the color structure tensor. Resulting transform appears to be a clear improvement of our previous work - that is the only proposition of color monogenic wavelets to our knowledge. This is shift and rotation invariant and efficiently represents multiscale color lines and edges.

Index Terms— Color Wavelets, Analytic, Monogenic, Wavelet transforms, Image analysis

1. INTRODUCTION

Because most signal processing tools are defined for scalar signals only; color images are most often handled by working either on intensity or with marginal schemes. Those methods are sometimes unable to consider some contours, and may also introduce false colors. Yet color information is fundamental in some applications such as medical imaging for instance.

Differential approaches are more flexible and allow analyzing vector functions of any dimension within the Riemannian manifolds framework [1]. Vector gradient and structure tensor provide some local variational analysis leading to efficient contour detection for example. Adaptive smoothing is also best performed with the color structure tensor [2]. Each color pixel is handled as an entity so the main visual content is well preserved *i.e.* no false color is introduced.

Signal processing approaches are not so developed for multidimensional signals. However recent improvements about 2D signals showed that the 2D *phase* and *frequency* concepts can lead to interesting local geometric analysis of grayscale images tied to a physical interpretation [3, 4]. Monogenic wavelets [5] perform multiscale local phase analysis. Unfortunately no color definition exists so far.

We propose to get the best of both worlds by injecting the color structure tensor formalism into monogenic wavelets. This paper provides a non-trivial color extension of monogenic signal and wavelets to perform coherent multiscale color 2D phase analysis. This definition is an improvement of our previous work [6].

Special Notations :

- Complex numbers : $z = a + \boldsymbol{j}b = |z|e^{\boldsymbol{j}\arg\{z\}}$
- 2D coordinate in bold : $\boldsymbol{x} = [x_1, x_2] \quad \boldsymbol{\omega} = [\omega_1, \omega_2]$
- Fourier transform : $s \overset{\mathcal{F}}{\longleftrightarrow} \hat{s} = \mathrm{TF}\{s\}$
- `max` : Maximum absolute value of a displayed dataset

This work is part of the French ANR project VERSO-CAIMAN.

2. MULTIRESOLUTION MONOGENIC ANALYSIS

This section recalls the classical *analytic signal* concept, the 2D monogenic signal and the associated monogenic wavelet proposed in [5]. These *signal processing* tools are well suited to describe scalar signals in terms of an *amplitude envelope* that is invariant under some transformations and a complementary *phase* data encoding the local shape of the signal. Further details will be found in [7, 4, 5].

2.1. Analytic signal

The analytic signal definition can be found in some signal processing manuals. Given a real signal $s(t)$, consider its Hilbert transform :

$$\{\mathcal{H}s\}(x) \overset{\mathcal{F}}{\longleftrightarrow} -\boldsymbol{j}\,\mathrm{sign}\,\omega\hat{s}(\omega) \tag{1}$$

The analytic signal $s_A(t)$ associated to $s(t)$ reads :

$$s_A(x) = s(x) + \boldsymbol{j}\{\mathcal{H}s\}(x) = A(x)e^{\boldsymbol{j}\varphi(x)} \tag{2}$$

The analytic signal is an AM/FM (Amplitude/frequency modulation) model and has mostly been used in communications. But another interpretation in terms of *local structure* is possible [3]. As illustrated Fig. 1, amplitude A traduces the relative presence of local energy while instant phase φ encodes some shape information. More precisely, $\varphi = 0$ or $\pm\pi$ indicates 'pulse'-like pieces of signal whereas $\varphi = \pm\frac{\pi}{2}$ indicates 'slopes'. This geometric interpretation motivated researchers to define a 2D extension to perform AM/FM image analysis. The most successful one is the monogenic signal.

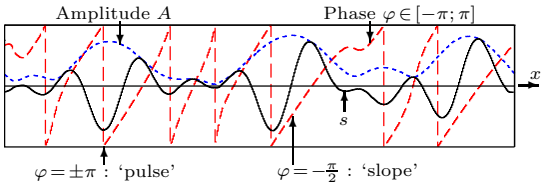

Fig. 1. Analytic signal associated to a scalar signal $s(x)$.

2.2. Monogenic signal

Given a 2D real (scalar) signal s, consider its Riesz transform :

$$\{\mathcal{R}s\}(\boldsymbol{x}) = s_R(\boldsymbol{x}) = s_{R_1} + \boldsymbol{j}s_{R_2} \overset{\mathcal{F}}{\longleftrightarrow} \frac{\omega_2 - \boldsymbol{j}\omega_1}{\|\boldsymbol{\omega}\|}\hat{s}(\boldsymbol{\omega}) \tag{3}$$

Note that this is the $\mathbb{C}$ embedding of the Riesz transform according to [8][1]. The monogenic signal s_M associated to s reads :

$$s_M = [s\ ;\ s_{R_1}\ ;\ s_{R_2}] = A[\cos\varphi\ ;\ \sin\varphi[\cos\theta\ ;\ \sin\theta]] \tag{4}$$

[1] This choice will help to define proper Riesz-based complex wavelets.

Where $\theta = \arg\{s_R\} \in [-\pi;\pi[$ is the local Riesz orientation along which a 1D Hilbert analysis is intrinsically done. This 1D analysis can be written :

$$s_A(\boldsymbol{x}) = s(\boldsymbol{x}) + \boldsymbol{j}|s_R(\boldsymbol{x})| = A(\boldsymbol{x})e^{\boldsymbol{j}\varphi(\boldsymbol{x})} \tag{5}$$

Amplitude of the monogenic signal traduces a local presence of some geometrical elements. The θ-phase gives the corresponding local orientation (we will see that it is equal to a gradient direction). The φ-phase results from an intrinsic Hilbert analysis along orientation θ. So the signal model here is an A-strong structure that is oriented along θ and looking like rather an *edge* ($\varphi \approx \pm\pi/2$) or a *line* ($\varphi \approx 0$ or π). Note that the good discrimination of lines and edges by local phase analysis was discussed in [9]. A recent extension - conformal monogenic signal [10] - also includes intrinsically 2D structures but this is out of the scope of this paper.

Since monogenic representation is clearly well suited to analyze *narrow-band* signals, it is natural to use it jointly with *subband* decomposition in order to handle any image.

2.3. Monogenic wavelet transform (MWT)

The scheme of [5] performs multiresolution monogenic analysis by using two parallel filterbanks. One 'primary' transform and a so-called 'Riesz-Laplace' wavelet transform. Multiresolution analyses are built around the nearly isotropic polyharmonic B-spline of [11] :

$$\beta_\gamma \overset{\mathcal{F}}{\longleftrightarrow} \left(\frac{4(\sin^2\frac{\omega_1}{2} + \sin^2\frac{\omega_2}{2}) - \frac{8}{3}\sin^2\frac{\omega_1}{2}\sin^2\frac{\omega_2}{2}}{\|\boldsymbol{\omega}\|^2} \right)^{\frac{\gamma}{2}} \tag{6}$$

which is a valid scaling function. The wavelet for the primary decomposition ψ is a Mexican hat-like nearly isotropic function and the 'Riesz-Laplace' wavelet ψ' is derived from it :

$$\psi(\boldsymbol{x}) = (-\Delta)^{\frac{\gamma}{2}}\ \beta_{2\gamma}(2\boldsymbol{x}) \qquad \psi' \overset{\mathcal{F}}{\longleftrightarrow} \frac{\boldsymbol{j}\omega_1 + \omega_2}{\|\boldsymbol{\omega}\|}\hat{\psi}(\boldsymbol{\omega}) \tag{7}$$

where the fractional Laplacian operator is defined by :

$$(-\Delta)^\alpha s \overset{\mathcal{F}}{\longleftrightarrow} \|\boldsymbol{\omega}\|^{2\alpha}\hat{s} \tag{8}$$

Let $\psi_{i,\boldsymbol{k}}(\boldsymbol{x}) = 2^i\psi(2^i\boldsymbol{x} - \boldsymbol{k}/2)$ be the i-scaled $\boldsymbol{k}$-shifted version of ψ (same for ψ'). It is shown [5] that $\psi_{i,\boldsymbol{k}}$ and $\psi'_{i,\boldsymbol{k}}$ form two wavelet frames ensuring perfect reconstruction and orthogonality across scales. In addition, their *operator-like behavior* induces :

$$\langle s, \psi_{i,\boldsymbol{k}}\rangle = (\psi_i * s)\,(2^{-(i+1)}\boldsymbol{k}) \tag{9}$$

$$\langle s, \psi'_{i,\boldsymbol{k}}\rangle = \mathcal{R}\,(\psi_i * s)\,(2^{-(i+1)}\boldsymbol{k}) \tag{10}$$

This means that wavelet coefficients form an exact monogenic signal at each scale. Both decompositions are merged into 3-vectors and turned into polar coordinate according to eq. (4) as illustrated Fig. 2. Interpretation of coefficients is the same as explained for the monogenic signal but now concerns different scales of the image. Contrary to most used wavelet transforms this one is non-separable to nearly achieve rotation invariance of the monogenic framework. Both wavelet frames correspond to a dyadic pyramid filterbank needing only one 2D wavelet for each transform [12]. Dyadic downsampling is done only at the low frequency branch leading to a total redundancy of 4:1.

This representation clearly fails to handle color images. Applying it on intensity of an image would induce a serious loss of information around ruptures of isoluminance. Using it marginally would have no more sense because independent geometric analyses would then be made on each color channel. In this case some processing such as thresholding would introduce false colors. We propose to extend the monogenic framework to color so that a coherent color orientation is considered and the 1D phase interpretation is kept.

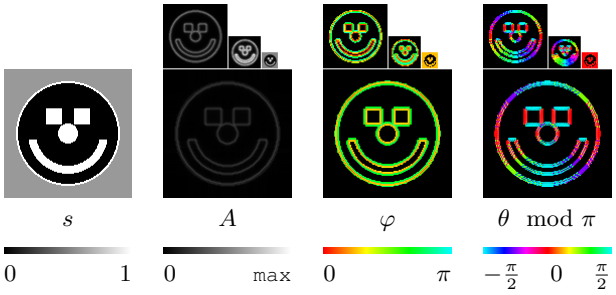

Fig. 2. MWT of image s. Orientation θ is shown modulo π for visual convenience. Phase values of small coefficients have no sense and are numerically unstable so they are replaced by black pixels.

3. EXTENDING TO COLOR IMAGES

This section aims at defining a sound color counterpart of the monogenic framework. Based on a link with the gradient and structure tensor features, we extract Riesz norm and direction from the well known color structure tensor formalism. The second step is to define the corresponding Hilbert analysis that completes the model.

3.1. Link between Riesz and the gradient

First let us recall definitions of the gradient

$$\nabla s = \left[\frac{\partial s}{\partial x}, \frac{\partial s}{\partial y}\right] = [s_x, s_y] \overset{\mathcal{F}}{\longleftrightarrow} [\boldsymbol{j}\omega_1\hat{s}, \boldsymbol{j}\omega_2\hat{s}] \tag{11}$$

and its associated structure tensor $T = h * [\nabla s]^T[\nabla s]$

$$T = \begin{bmatrix} h*s_x^2 & h*s_xs_y \\ h*s_xs_y & h*s_y^2 \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{12} & T_{22} \end{bmatrix} \tag{12}$$

tied to the smoothing kernel h. Eigenvalues and eigenvectors of T give information about direction and amplitude of the local maximum variation. Classical features are :

- Eigenvalues $\lambda_\pm = \frac{(T_{11}+T_{22}\pm\sqrt{(T_{22}-T_{11})^2+4T_{12}^2})}{2}$;
- Gradient norm $\mathcal{N} = \sqrt{\lambda_+ + \lambda_-}$
- Gradient direction $\theta = \frac{1}{2}\arctan\frac{2T_{12}}{T_{11}-T_{22}}$;

Now let us build a tensor formed with the Riesz components of s : $T_R = [s_{R_1}, s_{R_2}]^T[s_{R_1}, s_{R_2}]$. Construction from Unser *et. al.* [5] reveals a link between Riesz and gradient. Equations (3,8,11) give :

$$\mathcal{R}s = \left(-(-\Delta)^{-\frac{1}{2}}s_x\right) + \boldsymbol{j}\left(-(-\Delta)^{-\frac{1}{2}}s_y\right) \tag{13}$$

Using property $(-\Delta)^{\alpha_1}(-\Delta)^{\alpha_2} = (-\Delta)^{\alpha_1+\alpha_2}$ we get :

$$T_R = \begin{bmatrix} (-\Delta)^{-1}*s_x^2 & (-\Delta)^{-1}*s_xs_y \\ (-\Delta)^{-1}*s_xs_y & (-\Delta)^{-1}*s_y^2 \end{bmatrix} \tag{14}$$

The Riesz transform is equal to a classical structure tensor with particular smoothing kernel $(-\Delta)^{-1} \overset{\mathcal{F}}{\longleftrightarrow} 1/\|\boldsymbol{\omega}\|^2$. Finally, gradient norm $\mathcal{N}$ is equal to $|s_R|$ and Riesz orientation is equal to gradient direction. The particularity of the Riesz transform is its link with the monogenic framework allowing computation of local phase.

3.2. Color Riesz features

We now discuss how to retrieve Riesz features $\mathcal{N}$ and θ from a color image. The idea is to use the color structure tensor formalism.

The color structure tensor is the central tool of color differential approaches. The idea was first proposed by Di Zenzo in [13] but is now generally defined in Riemannian manifold embeddings [1] and used in general color imaging with PDE's and variational approaches [2]. Given a color image $s = (s^r, s^g, s^b)$, consider its marginal gradients along x and y $(s^r_x, s^g_x, s^b_x, s^r_y, s^g_y, s^b_y)$. The following color structure tensor is defined as follows :

$$T = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \quad \text{with} \quad \begin{array}{l} a = (s^r_x)^2 + (s^g_x)^2 + (s^b_x)^2 \\ b = 2(s^r_x s^r_y + s^g_x s^g_y + s^b_x s^b_y) \\ c = (s^r_y)^2 + (s^g_y)^2 + (s^b_y)^2 \end{array} \tag{15}$$

Norm and direction are $\mathcal{N} = a+c$ and $\theta = \frac{1}{2}\arctan(2b, a-c)$.

It is now trivial to derive the Riesz case since the color tensor is the sum of marginal tensors $T = T_r + T_g + T_b$. The Riesz transforms on red, green and blue channels can obviously be combined to form a smoothed version of T :

$$T^{\text{color}}_R = T^r_R + T^g_R + T^b_R = \begin{bmatrix} (-\Delta)^{-1}a & (-\Delta)^{-1}b \\ (-\Delta)^{-1}b & (-\Delta)^{-1}c \end{bmatrix} \tag{16}$$

The color counterpart of the Riesz transform is then $s^{\text{color}}_R = \mathcal{N}\, e^{\boldsymbol{j}\theta}$. The advantage of this generalization is that the structure tensor gives a coherent oriented analysis of *all* color ruptures. A 1D Hilbert analysis can now be made between s and $|s^{\text{color}}_R| = \mathcal{N}$ so we can adapt the monogenic framework to color images.

3.3. Color monogenic analysis

The key idea is to handle Euclidean norm $\|s\| = \sqrt{s_r^2 + s_g^2 + s_b^2}$. In the scalar case, this reduces to $|s|$. It is then interesting to rewrite the scalar model while separating absolute value and sign of s :

$$s = \underbrace{\sqrt{s^2 + \mathcal{N}^2}}_{A} \cos\big(\underbrace{\arg\{s + \boldsymbol{j}\mathcal{N}\}}_{\varphi \in [0;\pi[}\big) \tag{17}$$

$$= \underbrace{\sqrt{|s|^2 + \mathcal{N}^2}}_{A} \cos\big(\underbrace{\arg\{|s| + \boldsymbol{j}\mathcal{N}\}}_{\varphi_2 \in [0;\frac{\pi}{2}[}\big) \underbrace{s/|s|}_{\text{“sign”}} \tag{18}$$

The new phase φ_2 is a restricted version of φ to interval $[0; \frac{\pi}{2}]$. Fortunately we keep interesting qualitative information of edge/line discrimination. The vector counterpart relies on the separation of s into its Euclidean norm and its 'color axis' $\vec{u}$:

$$\vec{u} = s/\|s\| = [\cos\alpha\ ,\ \sin\alpha\cos\beta\ ,\ \sin\alpha\sin\beta] \tag{19}$$

The color monogenic model becomes :

$$s = \sqrt{\|s\|^2 + \mathcal{N}^2}\ \cos\big(\arg\{\|s\| + \boldsymbol{j}\mathcal{N}\}\big)\ \vec{u} \tag{20}$$

The main 4 components of the new color monogenic signal in Cartesian and spherical coordinate are therefore :

$$\begin{array}{ll} s^{\text{color}}_M = [\ s_r\ ,\ s_g\ ,\ s_b\ ,\ \mathcal{N}\ ] & \\ \text{Amplitude :} & A = \sqrt{s_r^2 + s_g^2 + s_b^2 + \mathcal{N}^2} \in [0; +\infty[\\ \text{1D Phase :} & \varphi_2 = \arg\{\|s\| + \boldsymbol{j}\mathcal{N}\} \in [0; \frac{\pi}{2}[\\ \text{Color axis :} & \left\{ \begin{array}{l} \alpha = \arg\{s_r + \boldsymbol{j}\sqrt{s_g^2 + s_b^2}\} \in [0; \pi[\\ \beta = \arg\{s_g + \boldsymbol{j}s_b\} \in [-\pi; \pi[\end{array} \right. \end{array} \tag{21}$$

Color Riesz orientation θ may be joined to the model but is useless for reconstructing original signal.

3.4. Color monogenic wavelets

The extension to the wavelet domain is trivial since the above construction relies on a marginal Riesz transform (non marginality occurs when combining marginal outputs into meaningful data). We define the two vector - color - wavelets :

$$\psi = [\psi\ ,\ \psi\ ,\ \psi] \qquad \psi' = [\psi'\ ,\ \psi'\ ,\ \psi'] \tag{22}$$

Obtained coefficients for scale i and position $\boldsymbol{k}$ are noted :

$$w_{i,\boldsymbol{k}} = [r_{i,\boldsymbol{k}}\ ,\ g_{i,\boldsymbol{k}}\ ,\ b_{i,\boldsymbol{k}}] \qquad w'_{i,\boldsymbol{k}} = [r'_{i,\boldsymbol{k}}\ ,\ g'_{i,\boldsymbol{k}}\ ,\ b'_{i,\boldsymbol{k}}] \tag{23}$$

And the monogenic signal is for each scale i :

$$w_M(\boldsymbol{k}) = [r_{i,\boldsymbol{k}}\ ,\ g_{i,\boldsymbol{k}}\ ,\ b_{i,\boldsymbol{k}}\ ,\ \|w'_{i,\boldsymbol{k}}\|] \tag{24}$$

Then amplitude, phase and color axis can be retrieved with eq. (21) and orientation with the main eigenvector of the color structure tensor formed by coefficients w' according to eq. (16). Let us now study this new transform from a practical viewpoint.

4. USING COLOR MONOGENIC WAVELETS

See examples of MWT on Fig. 3. Amplitude's invariance property is linked to a high visual coherence w.r.t. image contours. This is both due to the use of slightly redundant isotropic pyramids and to the invariance of the monogenic signal. Ruptures of isoluminance like the boundary between green and red big disks in first image are well detected thanks to the use of Euclidean norms of marginal outputs.

Coefficients of small amplitude encode uniform areas which need not any geometric analysis. Their phase are irrelevant and unused for reconstruction (Cartesian terms are all near-zero).

Phase φ_2 has a local monotone variation in the direction orthogonal to the contour. Its value at the position of maximum amplitude reveals the kind of rupture as explained section 2.2. This rich description of ruptures is exclusive to the phase concept [9] and so makes this signal processing approach competitive. In practice φ_2 is hard to visualize because of its fast variation w.r.t. resolution.

Ruptures occur from one color to another - drawing an axis in the color space. This axis is represented by the unit vector $\vec{u}$ that completes the phase data. In most cases $\vec{u}$ is slowly varying.

Additional directional analysis θ is equal to color gradient direction *i.e.* local color structure main spatial direction.

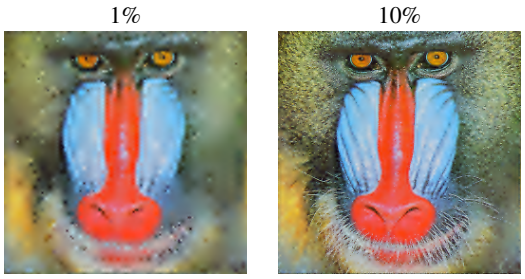

Fig. 4. Partial reconstructions with color MWT.

We show partial reconstructions Fig. 4. x percent of amplitude coefficients are kept while others are set to zero (discarding the phase). Each kept coefficient is complemented by the 3-angle phase data so it can be viewed like $3x\%$ rather than $x\%$. We can see that contours and colors are quite well preserved while textured areas are uniformly smoothed.

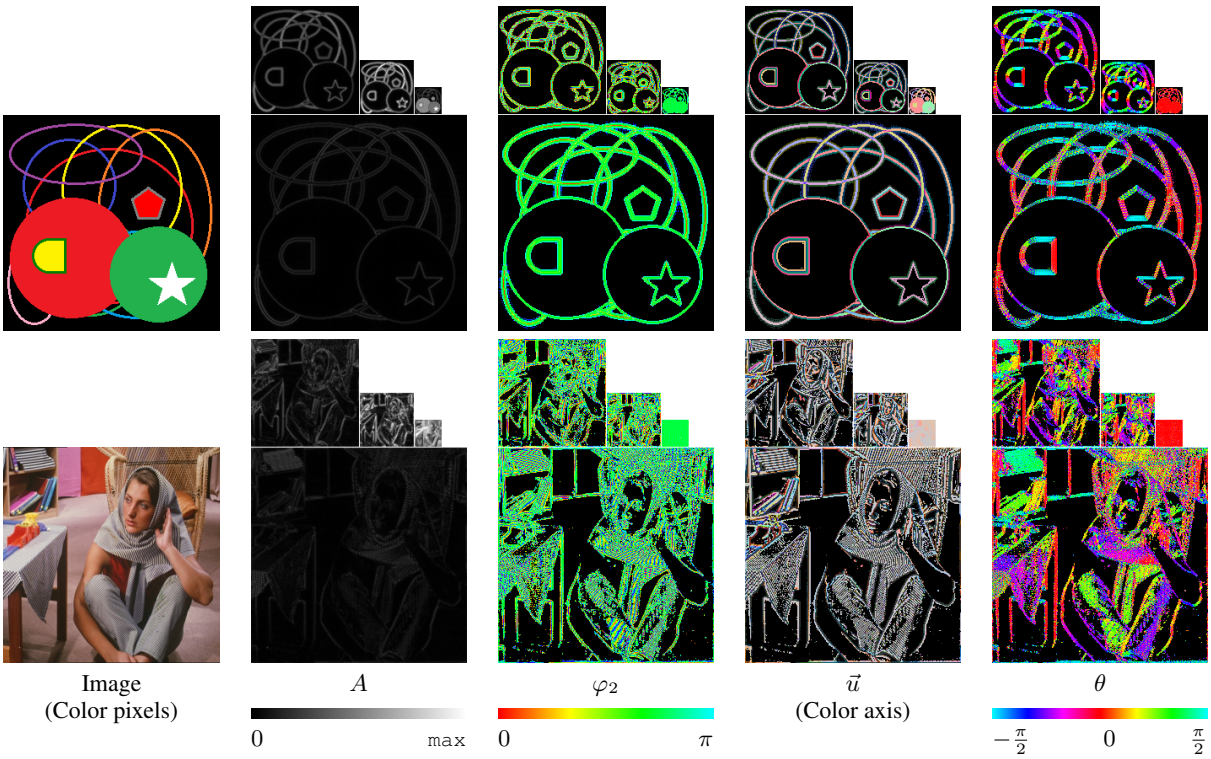

Fig. 3. Color monogenic wavelet transform on synthetic image and `barbara` image.

Finally, this monogenic wavelet representation of color images is consistent with the grayscale definition in terms of signal processing interpretation (phase and orientation) as well as it takes advantage of a well defined differential model for handling vector signals.

5. CONCLUSION

This paper presents a new proposition of color MWT by extending the grayscale MWT of Unser *et. al.* The definition is strongly linked to color structure tensor while offering a *signal processing* interpretation through the local phase concept. The transform is invariant by translation and rotation and gives coherent multiscale representation of color structures. This non marginal scheme avoid classical problem of false color and also detects all color contours.

The construction can easily be extended to the general multichannel case by considering general vector structure tensors. Our prospects include applications in texture analysis, study of a possible non redundant analytic representation as well as compression.

6. REFERENCES

[1] N. Sochen, R. Kimmel, and R. Malladi, "A general framework for low level vision," *IEEE Trans. Image Process.*, vol. 7, pp. 310–318, 1998.

[2] D. Tschumperlé and R. Deriche, "Vector-valued image regularization with pdes: A common framework for different applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 4, pp. 506–517, 2005.

[3] T. Bülow, "Hypercomplex spectral signal representation for the processing and analysis of images," *Thesis*, August 1999.

[4] M. Felsberg and G. Sommer, "The monogenic signal," *IEEE Trans. Signal Processing, vol. 49, no. 12, pp. 3136-3144*, 2001.

[5] M. Unser, D. Sage, and D. Van De Ville, "Multiresolution monogenic signal analysis using the Riesz-Laplace wavelet transform," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2402–2418, Nov. 2009.

[6] Raphaël Soulard and Philippe Carré, "Color monogenic wavelets for image analysis," in *IEEE ICIP*, Brussels, Belgium, Sept. 2011, pp. 277–280.

[7] D. Gabor, "Theory of communication," *JIEE*, vol. 93, no. 3, pp. 429–459, 1946.

[8] K. G. Larkin, D. Bone, and M. A. Oldfield, "Natural demodulation of two-dimensional fringe patterns: I. general background to the spiral phase quadrature transform," *J. Opt. Soc. Am.*, vol. 18 (8), pp. 1862–1870, 2001.

[9] P. Kovesi, "Edges are not just steps," in *ACCV*, Jan 2002, pp. pp 822–827.

[10] O. Fleischmann, L. Wietzke, and G. Sommer, "Image analysis by conformal embedding," *J. Math. Imaging Vis.*, vol. 40, no. 3, pp. 305–325, 2011.

[11] D. Van De Ville, T. Blu, and M. Unser, "Isotropic polyharmonic b-splines: Scaling functions and wavelets," *IEEE Trans. Image Process.*, vol. 14 :(11), pp. 1798–1813, 2005.

[12] M. Unser and D. Van De Ville, "The pairing of a wavelet basis with a mildly redundant analysis via subband regression," *IEEE Trans. Image Process.*, vol. 17, no. 11, pp. 1–13, Nov. 2008.

[13] S. Di Zenzo, "A note on the gradient of a multi-image," *Comput. Vis. Graph. Image Process.*, vol. 33, no. 1, pp. 116 – 125, 1986.